

Secure Broadcasting

Ashish Khisti, *Student Member, IEEE*, Aslan Tchamkerten,
and Gregory W. Wornell, *Fellow, IEEE*

Abstract

Wyner's wiretap channel is extended to parallel broadcast channels and fading channels with multiple receivers. In the first part of the paper, we consider the setup of parallel broadcast channels with one sender, multiple intended receivers, and one eavesdropper. We study the situations where the sender broadcasts either a common message or independent messages to the intended receivers. We derive upper and lower bounds on the common-message-secrecy capacity, which coincide when the users are reversely degraded. For the case of independent messages we establish the secrecy sum-capacity when the users are reversely degraded.

In the second part of the paper we apply our results to fading channels: perfect channel state information of all intended receivers is known globally, whereas the eavesdropper channel is known only to her. For the common message case, a somewhat surprising result is proven: a positive rate can be achieved independently of the number of intended receivers. For independent messages, an *opportunistic* transmission scheme is presented that achieves the secrecy sum-capacity in the limit of large number of receivers. Our results are stated for a fast fading channel model. Extensions to the block fading model are also discussed.

Index Terms

Wiretap channel, information theoretic secrecy, confidential messages, parallel channels, fading channels, multiuser diversity, multicasting

I. INTRODUCTION

A number of emerging applications require a “key distribution mechanism” to selectively broadcast confidential messages to intended receivers. For example in *pay TV systems*, a content provider wishes to selectively broadcast a certain program to a subset of customers who have subscribed to it. An online key distribution mechanism would allow the service provider to distribute a decryption key to these intended receivers while securing it from potential eavesdroppers. The program could then be encrypted via standard cryptographic protocols, so that only users who have access to the decryption key could view it. Indeed, in the absence of such a mechanism, current solutions rely on variants of traditional public key cryptography (see, e.g., [5]) and are vulnerable to attacks such as piracy [7].

An information theoretic framework for perfect secrecy was developed by Shannon [18], and the problem of broadcasting confidential messages was originally formulated by Wyner [22]. Wyner considered a special broadcast channel (also known as the wiretap channel): one sender, an intended receiver, and one eavesdropper. He characterized the tradeoff between the rate to the intended receiver and the equivocation at the eavesdropper when the eavesdropper has a degraded channel compared to the intended receiver. This formulation has been generalized for non-degraded broadcast channels in [3], and applied to Gaussian channels in [13].

While the results for wire-tap channels are rather surprising in that they show that it is possible to achieve a positive rate while keeping the eavesdropper in near-perfect equivocation, they also provide some disappointing facts for degraded channels, such as Gaussian [13]. First, the secrecy capacity is positive only if the eavesdropper is noisier than the intended receiver. This may not be the case in practice. Second, in the limit of high signal-to-noise ratio (SNR), the secrecy capacity approaches a constant and does not exhibit a logarithmic growth with power. Thus, physical layer secrecy comes at a price in throughput

This work was supported in part by NSF under Grant No. CCF-0515109.

The authors are with the Massachusetts Institute of Technology. Email: {khisti,tcham,gww}@mit.edu. This work was presented in part at the 44th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, September 26-29, 2006.

and this may have prompted many practical cryptographic solutions to be based upon other notions of security, such as computational security [5]. Note, however, that such solutions require an off-line key distribution mechanism which may not be practical in emerging applications.

The wiretap channel has received renewed interest in some recent works that consider a wireless environment. There the eavesdropper is not always stronger than the intended receivers due to time variations in channel gains. These variations in turn can be exploited to communicate securely by transmitting to the receivers that have a strong channel. Such coding strategies may yield a practical approach for secure communication without an off-line key agreement.

In the present work we extend Wyner's wiretap channel to parallel broadcast channels with one sender, multiple intended receivers, and one eavesdropper. We consider two situations: all intended receivers get a common message or independent messages. We first derive upper and lower bounds on the common-message-secrecy-capacity. These bounds coincide when the users are reversely degraded. Perhaps the main observation is that, to achieve the common message capacity, independent codebooks are used on each parallel channel, and each receiver jointly decodes its received sequences. Next, we consider the case where the intended receivers get independent messages. We establish the secrecy capacity for the reversely degraded case. The achievable scheme is simple: transmit to the strongest user on each parallel channel and use independent codebooks across the channels. Our results for the parallel broadcast channels can be viewed as generalizations of the results in [6] which considers a similar setup without the presence of an eavesdropper.

Our study on parallel channels provides insights to the problem of broadcasting confidential messages over fading channels. In the second part of the paper we consider an i.i.d. fading model. We assume the intended receivers' channel state information (CSI) is revealed to all communicating parties (including the eavesdropper), while the eavesdropper's channel gains are revealed only to her.

We first examine the case when a common message needs to be delivered to all intended receivers in the presence of potential eavesdroppers. We refer to this problem as secure multicasting. We present a scheme that exploits CSI at the transmitter and achieves a rate that does not decay to zero with increasing number of receivers. Note that, without a secrecy constraint, transmitter CSI appears to be of little value for multicasting over ergodic channels. Indeed the capacity appears to be not too far from the maximum achievable rate with a flat power allocation scheme. The secrecy constraint adds a new twist to the multicasting problem as it requires to consider protocols that exploit transmitter CSI.

For the case of independent messages, we consider an opportunistic scheme that selects the user with the strongest channel at each time. We use Gaussian wiretap codebooks for each intended receiver and show that this scheme achieves the sum capacity in the limit of large number of receivers. Our results can be interpreted as the wiretap analog of the multiuser diversity results in settings without secrecy constraint (see, e.g., [20]).

In related works, the Gaussian wiretap channel was extended to parallel channels in [23]. More recently, the case of discrete memoryless parallel channels with one receiver and one eavesdropper has been studied in [15], [16]. The wiretap setting has been also studied for fading channels in [1], [10], [17]. All these works consider the setup of one sender, one receiver, and one eavesdropper.

There is also a vast literature on multiuser diversity in broadcast channels with independent messages starting from the results in [14], [19]. However, to the best of our knowledge, the present work is the first to consider the impact of multiuser diversity on secrecy systems. As discussed before, the case of a common message has received much less attention in the literature. The problem of transmitting a common message on parallel channels has been studied in [6], [11] but we are not aware of a general treatment of this problem for fading channels (without an eavesdropper). We hope that the secrecy constraint creates renewed interest in the study of common message broadcast problems, given its application to key distribution.

We use the following notation. Upper case letters are used for random variables and the lower case for their realizations. The notation s^n denotes a vector of length n . Vector quantities related to the eavesdropper have a subscript e , e.g., y_e^n , while the ones of the intended receivers are subscripted by the user number, e.g., y_i^n . We use the subscript i to index the receivers and the subscript j to index the channels. We use

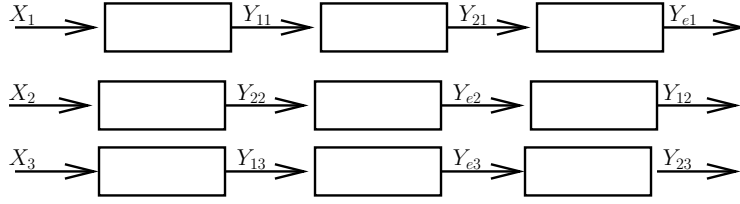


Fig. 1. An example of reversely degraded parallel channel in Definition 2 with one sender, $K = 2$ users, one eavesdropper, and $M = 3$ channels.

the letter t to denote the discrete time index. If there is an ordering of users on a given channel, the strongest user on channel j will be denoted by π_j . The set of ordered users on channel j is denoted as $\pi_j(1), \pi_j(2), \dots$. We use the notation $p(X_j)$ to denote the probability mass function of random variable X_j .

II. PARALLEL CHANNELS: MODEL

In our setup, there are M parallel channels for communication, one sender, K intended receivers, and one eavesdropper.

Definition 1 (Product Broadcast Channel): An (M, K) product broadcast channel consists of one sender, K receivers, one eavesdropper, and M channels. The channels have finite input and output alphabets, are memoryless and independent of each other, and are characterized by their transition probabilities given by

$$\Pr(\{y_{1j}^n, y_{2j}^n, \dots, y_{Kj}^n, y_{ej}^n\}_{j=1, \dots, M} \mid \{x_j^n\}_{j=1, \dots, M}) = \prod_{j=1}^M \prod_{t=1}^n \Pr(y_{1j}(t), y_{2j}(t), \dots, y_{Kj}(t), y_{ej}(t) \mid x_j(t)) \quad (1)$$

for $j = 1, 2, \dots, M$, where $x_j^n = x_j(1), x_j(2), \dots, x_j(n)$ denotes the sequence of symbols transmitted on channel j , and where $y_{ij}^n = y_{ij}(1), y_{ij}(2), \dots, y_{ij}(n)$ denotes the sequence of symbols received by user i on channel j from time 1 up to n . The alphabets of the X_j 's and Y_{ij} 's are denoted by \mathcal{X} and \mathcal{Y} respectively.

Of particular interest is a special class of reversely degraded broadcast channels.

Definition 2 (Reversely Degraded Broadcast Channel): An (M, K) reversely degraded broadcast channel is an (M, K) product broadcast channel, where each of the M parallel channels is degraded in a certain order. For some permutation $\pi_j(1), \pi_j(2), \dots, \pi_j(K+1)$ of the set $\{1, 2, \dots, K, e\}$ of the $K+1$ -receivers, a Markov chain $X_j \rightarrow Y_{\pi_j(1)} \rightarrow Y_{\pi_j(2)} \rightarrow \dots \rightarrow Y_{\pi_j(K+1)}$ can be specified.

Remark 1: Note that in Definition 2 the order of degradation can be different across the channels, so the overall channel may not be degraded. An example of reversely degraded parallel channel is shown in Fig 1. Also, on any parallel channels component, the K users and the eavesdropper are *physically* degraded. Our capacity results will, however, only depend on the marginal distribution of receivers on each channel (see Fact 1 below). Accordingly, these results also hold for a larger class of channels where receivers on each channel are stochastically degraded.

III. PARALLEL CHANNELS: COMMON MESSAGE

In this section we consider the case where all the receivers are interested in only a common message. This common message must be protected from the eavesdropper in the sense described below.

Definition 3: A $(n, 2^{nR})$ code consists of a message set $\mathcal{W} = \{1, 2, \dots, 2^{nR}\}$, a (possibly stochastic) mapping $\omega_n : \mathcal{W} \rightarrow \underbrace{\mathcal{X}^n \times \mathcal{X}^n \times \dots \times \mathcal{X}^n}_{M \text{ times}}$ from the message set to the codewords for the M channels, and a decoder $\Phi_{i,n} : \underbrace{\mathcal{Y}^n \times \mathcal{Y}^n \times \dots \times \mathcal{Y}^n}_{M \text{ times}} \rightarrow \mathcal{W}$ for $i = 1, 2, \dots, K$ at each receiver. We denote the

message estimate at decoder i by \hat{W}_i . A common-message-secrecy-rate R is achievable if, for any $\varepsilon > 0$, there exists a length n code such that $\Pr(W \neq \hat{W}_i) \leq \varepsilon$ for $i = 1, 2, \dots, K$, while

$$\frac{1}{n}H(W|Y_{e1}^n, Y_{e2}^n, \dots, Y_{eK}^n) \geq R - \varepsilon. \quad (2)$$

The common-message-secrecy-capacity is the supremum over all achievable rates.

Remark 2: Wyner's formulation considers the rate-equivocation region (R, R_e) with $\frac{1}{n}H(W) \geq R$ and $\frac{1}{n}H(W|Y_e^n) \geq R_e$. The secrecy-capacity constitutes the special case when $R = R_e$. In the key-distribution application of interest, the key length is limited by the equivocation rate R_e — the minimum number of bits the eavesdropper needs to guess to decode the message. Accordingly, the secrecy capacity is of primary interest.

A. Main Results

Our main result is the characterization of upper and lower bounds on the common-message-secrecy-capacity for the product channel model (1). The bounds coincide for the reversely degraded model.

To state our upper bound we introduce the following additional notation. For any $j = 1, 2, \dots, M$, let \mathcal{P}_j denote the collection of all joint distributions $p'(Y_{1j}, Y_{2j}, \dots, Y_{Kj}, Y_{ej}|X_j)$ with the same marginal distribution as $p(Y_{1j}|X_j), p(Y_{2j}|X_j), \dots, p(Y_{Kj}|X_j), p(Y_{ej}|X_j)$. Let $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \times \dots \times \mathcal{P}_M$ denote the cartesian product of these sets across the channels.

Lemma 1 (Upper Bound): For the product broadcast channel model in Definition 1, an upper bound on the secrecy capacity is given by

$$R_{K,M}^{+, \text{common}} \triangleq \min_{\mathcal{P}} \max_{\prod_{j=1}^M p(X_j)} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M I(X_j; Y_{ij}|Y_{ej}) \quad (3)$$

where the first minimum is over all the joint distributions

$$\{p'(Y_{1j}, Y_{2j}, \dots, Y_{Kj}, Y_{ej}|X_j)\}_{j=1}^M \in \mathcal{P}.$$

Lemma 2 (Lower Bound): An achievable common-message-secrecy-rate for the product broadcast channel model Definition 1 is¹

$$R_{K,M}^{-, \text{common}} \triangleq \max_{\substack{\prod_{j=1}^M p(U_j) \\ \{X_j = f_j(U_j)\}_{j=1, \dots, M}}} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M \{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+. \quad (4)$$

The random variables U_1, U_2, \dots, U_M are independent over some alphabet \mathcal{U} , and each $f_j : \mathcal{U} \rightarrow \mathcal{X}$, $j = 1, \dots, M$ is a (possibly stochastic²) mapping from the \mathcal{U} to \mathcal{X} .

Our upper and lower bounds coincide for the case of reversely degraded product channels.

Theorem 1: The common-message-secrecy-capacity for the reversely degraded channel model in Definition 2 is given by

$$C_{K,M}^{\text{common}} = \max_{\prod_{j=1}^M p(X_j)} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M I(X_j; Y_{ij}|Y_{ej}). \quad (5)$$

Note that the expression in (5) is evaluated for the joint distribution induced by the reversely degraded channel. This distribution is the worst-case distribution in the set \mathcal{P} in (3).

Remark 3: Our achievable rate expression in (4) involves optimization over the auxiliary random variables U_j and the stochastic mappings $f_j(\cdot)$. As noted in [3], the expression $I(U_j; Y_{ij}) - I(U_j; Y_{ej})$ is in general not convex in $p(X_j|U_j)$, hence the optimal $f_j(\cdot)$ need not be deterministic functions. However, for the special reversely degraded case in Theorem 1, the choice $X_j = U_j$ is optimal (see Section III-D).

¹ $\{v\}^+$ stands for $\max\{0, v\}$.

²For each $u \in \mathcal{U}_j$, a stochastic mapping $f_j : \mathcal{U} \rightarrow \mathcal{X}$ produces a random element in \mathcal{X} .

The proof of the upper bound in Lemma 1 is a rather straightforward extension of Wyner's converse for the single user wiretap channel. The achievability proof in Lemma 2 is more interesting. When specialized to the case of no eavesdropper, it provides a different capacity achieving scheme than the one considered in [6].

B. Upper Bound

Fact 1: The common-message-secrecy-capacity for the wiretap channel depends only on the marginal distributions $p(Y_{1j}|X_j), p(Y_{2j}|X_j), \dots, p(Y_{Kj}|X_j)$ in (1) and not on the joint distribution $p(Y_{1j}, Y_{2j}, \dots, Y_{Kj}|X_j)$ for each $j = 1, 2, \dots, M$.

The proof of this fact is essentially the same as the proof for broadcast channels without secrecy constraint (see, e.g., [3]).

The following property will be used in the upper bound derivation but, also, in other subsequent proofs.

Fact 2: For any random variables X , Y , and Z the quantity $I(X; Y|Z)$ is concave in $p(X)$. The proof is implicit in the arguments in [21]. We provide it in Appendix I for completeness.

Suppose there exists a sequence of $(n, 2^{nR})$ codes such that, for every $\varepsilon > 0$, as $n \rightarrow \infty$

$$\begin{aligned} \Pr(W \neq \hat{W}_i) &\leq \varepsilon, \quad i = 1, 2, \dots, K \\ \frac{1}{n} I(W; Y_{e1}^n, \dots, Y_{eM}^n) &\leq \varepsilon. \end{aligned} \quad (6)$$

We first note that from Fano's Lemma we have

$$\frac{1}{n} H(W|Y_{i1}^n, Y_{i2}^n, \dots, Y_{iM}^n) \leq \frac{1}{n} + \varepsilon R \quad i = 1, 2, \dots, K. \quad (7)$$

Combining (6) and (7) we have, for all $i = 1, 2, \dots, K$ and $\varepsilon' = \varepsilon + \frac{1}{n} + \varepsilon R$,

$$\begin{aligned} nR &\leq I(W; Y_{i1}^n, \dots, Y_{iM}^n) - I(W; Y_{e1}^n, \dots, Y_{eM}^n) + n\varepsilon' \\ &\leq I(W; Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n) + n\varepsilon' \\ &= h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n) - h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n, W) \\ &\leq h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n) - h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n, X_1^n, \dots, X_M^n, W) \\ &= h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n) - h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n, X_1^n, \dots, X_M^n) \end{aligned} \quad (8)$$

$$= h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n) - \sum_{j=1}^M h(Y_{ij}^n | X_j^n, Y_{ej}^n) + n\varepsilon' \quad (9)$$

$$\begin{aligned} &\leq \sum_{j=1}^M h(Y_{ij}^n | Y_{ej}^n) - \sum_{j=1}^M h(Y_{ij}^n | X_j^n, Y_{ej}^n) + n\varepsilon' \\ &\leq \sum_{j=1}^M I(X_j^n; Y_{ij}^n | Y_{ej}^n) + n\varepsilon', \end{aligned} \quad (10)$$

where (8) follows from the fact that $W \rightarrow (X_1^n, \dots, X_M^n, Y_{e1}^n, \dots, Y_{eM}^n) \rightarrow (Y_{i1}^n, \dots, Y_{iM}^n)$ form a Markov chain, and (9) holds because the parallel channels are mutually independent in (1) so that

$$h(Y_{i1}^n, \dots, Y_{iM}^n | Y_{e1}^n, \dots, Y_{eM}^n, X_1^n, \dots, X_M^n) = \sum_{j=1}^M h(Y_{ij}^n | X_j^n, Y_{ej}^n).$$

We now upper bound each term in the summation (10). We have

$$I(X_j^n; Y_{ij}^n | Y_{ej}^n) \leq \sum_{k=1}^n I(X_j(k); Y_{ij}(k) | Y_{ej}(k)) \quad (11)$$

$$= \sum_{k=1}^n I(X_j(k); Y_{ij}(k), Y_{ej}(k)) - I(X_j(k); Y_{ej}(k)) \quad (12)$$

$$= nI(X_j; Y_{ij}, Y_{ej} | Q) - nI(X_j; Y_{ej} | Q) \quad (13)$$

$$= nI(X_j; Y_{ij} | Y_{ej}, Q) \leq nI(X_j; Y_{ij} | Y_{ej}), \quad (14)$$

where (11) follows from the fact that the channel is memoryless, and (13) is obtained by defining Q to be a (time-sharing) random variable uniformly distributed over $\{1, 2, \dots, n\}$ independent of everything else. The random variables (X_j, Y_{ij}, Y_{ej}) are such that, conditioned on $Q = k$, they have the same joint distribution as $(X_j(k), Y_{ij}(k), Y_{ej}(k))$. Finally (14) follows from the fact that the mutual information is concave with respect to the input distribution $p(X_j)$ as stated in Fact 2.

Combining (14) and (9) we have

$$\begin{aligned} R &\leq \sum_{j=1}^M I(X_j; Y_{ij} | Y_{ej}) + \varepsilon', \quad i = 1, 2, \dots, K \\ &= \min_{1 \leq i \leq K} \sum_{j=1}^M I(X_j; Y_{ij} | Y_{ej}) + \varepsilon' \end{aligned} \quad (15)$$

$$\leq \max_{\prod_{j=1}^M p(X_j)} \min_{1 \leq i \leq K} \sum_{j=1}^M I(X_j; Y_{ij} | Y_{ej}) + \varepsilon'. \quad (16)$$

The above bound (16) depends on the joint distribution across the channels. Accordingly, we tighten the upper bound by considering the worst distribution in $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \times \dots \times \mathcal{P}_M$ which gives

$$R \leq \min_{\mathcal{P}} \max_{\prod_{j=1}^M p(X_j)} \min_{1 \leq i \leq K} \sum_{j=1}^M I(X_j; Y_{ij} | Y_{ej}) + \varepsilon'. \quad (17)$$

C. Lower Bound

We first informally present the main ideas in our achievability scheme. We construct M independent codebooks, one for each channel, denoted as $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_M$. The structure of the codebooks is shown in Fig. 2. Each \mathcal{C}_j has $2^{n(R+I(U_j; Y_{ej}))}$ codewords, randomly partitioned into 2^{nR} *message bins* — there are $2^{nI(U_j; Y_{ej})}$ codewords per bin. Given a message W , the encoder selects M codewords as follows. On channel j , it looks into the bin corresponding to message W in \mathcal{C}_j and randomly selects a codeword in this bin. Each intended receiver attempts to find a message that is jointly typical with its received sequences. An appropriate choice of R guarantees successful decoding with high probability for each intended receiver, and near perfect equivocation at the eavesdropper.

We now provide a formal description of our coding scheme.

Fix the distributions $p(U_1), p(U_2), \dots, p(U_M)$ and the (possibly stochastic) functions $f_1(\cdot), \dots, f_M(\cdot)$. Let ε_E and ε_R be positive constants, to be quantified later. With respect to these quantities, define

$$\begin{aligned} R &= \min_{1 \leq i \leq K} \sum_{j=1}^M \{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+ - \varepsilon_R \\ R_{ej} &= I(U_j; Y_{ej}) - \varepsilon_F, \quad j = 1, 2, \dots, M. \end{aligned} \quad (18)$$

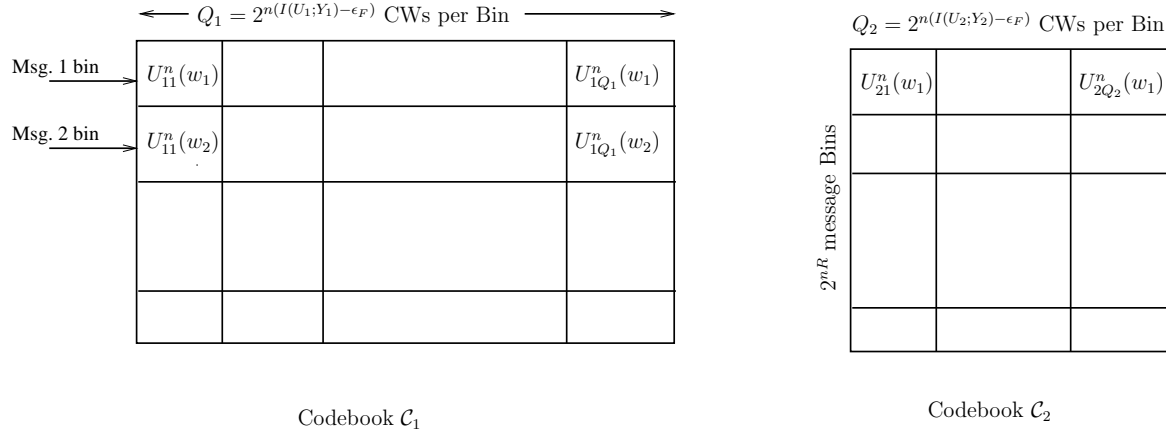


Fig. 2. Structure of the codebooks in our coding scheme for the case of two parallel channels. Each codebook has 2^{nR} message bins and $Q_j \approx 2^{n(I(U_j;Y_{ej})-\epsilon_F)}$ codewords per message bin. Thus the size of bins depends on the mutual information of the eavesdropper on the corresponding channel. This flexible binning enables to confuse the eavesdropper on each channel. Note that \mathcal{C}_1 and \mathcal{C}_2 above have the same number of rows but different number of columns. The codewords for message w_k in \mathcal{C}_j are labeled as $u_{j1}^n(w_k), \dots, u_{jQ_j}^n(w_k)$.

In what follows, whenever typicality is mentioned it is intended to be ε -weak typicality (see, e.g., [2]). The set $T(U_j)$ denotes the set of all sequences that are typical with respect to distribution $p(U_j)$ and the set $T(X_j, U_j)$ denotes the set of all jointly typical sequences (x_j^n, u_j^n) with respect to the distribution $p(X_j, U_j)$. $T_{u_j^n}(X_j|U_j)$ denotes the set of all sequences x_j^n conditionally typical with respect to a given sequence u_j^n according to $p(X_j|U_j)$.

1) Codebook Generation:

- Codebook \mathcal{C}_j for $j = 1, 2, \dots, M$ has a total of $M_j = 2^{n(R+R_{ej})}$ length n codeword sequences. Each sequence is selected uniformly and independently from the set $T(U_j)$.
- We randomly partition the M_j sequences into 2^{nR} message bins so that there are $Q_j = 2^{nR_{ej}}$ codewords per bin.
- The set of codewords associated with bin w in codebook \mathcal{C}_j is denoted as

$$\mathcal{C}_j(w) = \{u_{j1}^n(w), u_{j2}^n(w), \dots, u_{jQ_j}^n(w)\}, \quad w = 1, 2, \dots, 2^{nR}, \quad j = 1, 2, \dots, M. \quad (19)$$

Note that $\mathcal{C}_j = \bigcup_{w=1}^{2^{nR}} \mathcal{C}_j(w)$ is the codebook on channel j .

2) *Encoding*: To encode message w , the encoder randomly and uniformly selects a codeword in the set $\mathcal{C}_j(w)$ for all $1 \leq j \leq M$. Specifically,

- Select M integers k_1, k_2, \dots, k_M , where k_j is selected independently and uniformly from the set $\{1, 2, \dots, Q_j\}$.
- Given a message w , select a codeword $u_{jk_j}^n(w)$ from codebook $\mathcal{C}_j(w)$ for $j = 1, 2, \dots, M$.
- The transmitted sequence on channel j is denoted by $x_j^n = x_j(1), x_j(2), \dots, x_j(n)$. The symbol $x_j(t)$ is obtained by applying the (possibly stochastic) function $f_j(\cdot)$ on the t^{th} element of the codeword $u_{jk_j}^n(w)$.

3) *Decoding*: Receiver i , based on its observations $(y_{i1}^n, y_{i2}^n, \dots, y_{iM}^n)$ from the M parallel channels, declares message w according to the following rule.

- Let $\mathcal{S}_i = \{j | 1 \leq j \leq M, I(U_j; Y_{ij}) > I(U_j; Y_{ej})\}$ denote the set of channels where receiver i has larger mutual information than the eavesdropper. The receiver only considers the outputs y_{ij}^n from these channels.
- Receiver i searches for a message w such that, for each $j \in \mathcal{S}_i$, there is an index l_j such that $(u_{jl_j}^n(w), y_{ij}^n) \in T(U_j, Y_{ij})$. If a unique w has this property, the receiver declares it as the transmitted message. Otherwise, the receiver declares an arbitrary message.

4) *Error Probability:* We show that, averaged over the ensemble of codebooks, the error probability is smaller than a constant ε' (to be specified), which approaches zero as $n \rightarrow \infty$. This demonstrates the existence of a codebook with error probability less than ε' . We do the analysis for user i and, without loss of generality, assume that message w_1 is transmitted.

- **False Reject Event:** Let \mathcal{E}_{1j}^c be the event $\{(U_{jk_j}^n(w_1), Y_{ij}^n) \notin T(U_j, Y_{ij})\}$. Since $U_j^n \in T(U_j)$ by construction and Y_{ij} is obtained by passing U_j through a DMC, it follows that $\Pr(\mathcal{E}_{1j}^c) \leq \delta$, where $\delta \rightarrow 0$ as $\varepsilon \rightarrow 0$. Accordingly if \mathcal{E}_1^c denotes the event that message w_1 does not appear typical, then we have

$$\Pr(\mathcal{E}_1^c) = \Pr\left(\bigcup_{j=1}^M \mathcal{E}_{1j}^c\right) \leq M\delta. \quad (20)$$

- **False Accept Event:** As before, let $\mathcal{S}_i \subseteq \{1, 2, \dots, M\}$ denote the subset of channels for which $I(U_j; Y_{ij}) > I(U_j; Y_{ej})$. In what follows the index j will only refer to channels in \mathcal{S}_i . Let \mathcal{E}_{rj} denote the event that there is a codeword in the set $\mathcal{C}_j(w_r)$ ($r > 1$) typical with Y_{ij}^n . Also let \mathcal{E}_r be the event that message w_r has a codeword typical on every channel.

$$\begin{aligned} \Pr(\mathcal{E}_{rj}) &= \Pr(\exists l \in \{1, 2, \dots, Q_j\} : (U_{jl}^n(w_r), Y_{ij}^n) \in T(U_j, Y_{ij})), \quad j \in \mathcal{S} \\ &\leq \sum_{l=1}^{Q_j} \Pr((U_{jl}^n(w_r), Y_{ij}^n) \in T(U_j, Y_{ij})) \\ &\leq \sum_{l=1}^{Q_j} 2^{-n(I(U_j; Y_{ij}) - 3\delta)} \\ &\leq 2^{-n(I(U_j; Y_{ij}) - I(U_j; Y_{ej}) - 3\delta + \varepsilon_F)}, \end{aligned}$$

where the last inequality follows since $Q_j = 2^{n(I(U_j; Y_{ej}) - \varepsilon_F)}$. Finally, the probability of \mathcal{E}_r can be computed as

$$\begin{aligned} \Pr(\mathcal{E}_r) &= \Pr\left(\bigcap_{j \in \mathcal{S}_i} \mathcal{E}_{rj}\right) \\ &= \prod_{j \in \mathcal{S}_i} \Pr(\mathcal{E}_{rj}) \\ &= 2^{-n \sum_{j \in \mathcal{S}_i} (I(U_j; Y_{ij}) - I(U_j; Y_{ej}) - 3\varepsilon + \varepsilon_F)} \\ &= 2^{-n \sum_{j=1}^M (\{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+ - 3\varepsilon + \varepsilon_F)}, \end{aligned} \quad (21)$$

where (21) follows by independence of codebooks and channels. The probability of false accept event \mathcal{E}_F is then given by

$$\begin{aligned} \Pr(\mathcal{E}_F) &= \Pr\left(\bigcup_{r=2}^{2^{nR}} \mathcal{E}_r\right) \\ &\leq 2^{nR} 2^{-n \sum_{j=1}^M (\{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+ - 3\delta + \varepsilon_F)} \\ &\leq 2^{-n(3M\delta - M\varepsilon_F + \varepsilon_R)}, \end{aligned}$$

which vanishes with increasing n for any ε_R and ε_F that satisfy the relation $\varepsilon_R > M\varepsilon_F - 3M\delta > 0$. The probability of error averaged over the ensemble of codebooks is less than $\varepsilon' = \max(M\delta, 2^{-n(3M\delta - M\varepsilon_F + \varepsilon_R)})$. This demonstrates the existence of a codebook with error probability less than ε' .

5) *Secrecy Analysis*: We now bound the equivocation at the eavesdropper for a typical code in the ensemble. Informally, since the codebook \mathcal{C}_j has $2^{n(I(U_j; Y_{ej}) - \varepsilon_F)}$ codewords per bin, the eavesdropper's equivocation is near perfect when observing the output of channel j , i.e., $\frac{1}{n}I(W; Y_{ej}^n) \leq \varepsilon'_F$ for some ε'_F (to be specified) such that $\varepsilon'_F \rightarrow 0$ as $\varepsilon_F \rightarrow 0$. Since we are sending the same message on each of the M channels, the eavesdropper can potentially reduce the equivocation by combining the channel outputs. However in doing so, his equivocation reduces by at most $M\varepsilon'_F$ since the codewords on each channel are independently selected.³

The following Lemma is proved in Appendix II.

Lemma 3: A typical code from the ensemble in our achievability scheme satisfies the following: For any $j = 1, 2, \dots, M$, we have $\frac{1}{n}I(W; Y_{ej}^n) \leq \varepsilon'_F$, where $\varepsilon'_F = \varepsilon'_F(\delta, \varepsilon_F)$ tends to zero as $\delta \rightarrow 0$ and $\varepsilon_F \rightarrow 0$.

Using the above lemma we now upper bound the mutual information at the eavesdropper as

$$\frac{1}{n}I(W; Y_{e1}^n, \dots, Y_{eM}^n) = h(Y_{e1}^n, \dots, Y_{eM}^n) - h(Y_{e1}^n, \dots, Y_{eM}^n | W) \quad (22)$$

$$= h(Y_{e1}^n, \dots, Y_{eM}^n) - \sum_{j=1}^M h(Y_{ej}^n | W) \quad (23)$$

$$\leq \sum_{j=1}^M I(W; Y_{ej}^n) \leq Mn\varepsilon'_F, \quad (24)$$

where $h(Y_{e1}^n, \dots, Y_{eM}^n | W) = \sum_{j=1}^M h(Y_{ej}^n | W)$ since the codewords in the sets $\mathcal{C}_1(W), \mathcal{C}_2(W), \dots, \mathcal{C}_M(W)$ are independently selected.

Hence the normalized mutual information increases only by a fixed amount due to observations on multiple channels. By choosing ε in (2) to equal $M\varepsilon'_F$, we satisfy the secrecy constraint.

D. Capacity Result of Theorem 1

The result of Theorem 1 follows directly from Lemma 1 and 2. For the reversely degraded broadcast channel we have, for all i and j , that either $X_j \rightarrow Y_{ij} \rightarrow Y_{ej}$ or $X_j \rightarrow Y_{ej} \rightarrow Y_{ij}$ holds. If $X_j \rightarrow Y_{ij} \rightarrow Y_{ej}$ holds, then

$$\begin{aligned} I(X_j; Y_{ij}) - I(X_j; Y_{ej}) &= I(X_j; Y_{ij}, Y_{ej}) - I(X_j; Y_{ej}) \\ &= I(X_j; Y_{ij} | Y_{ej}). \end{aligned}$$

Instead, if $X_j \rightarrow Y_{ej} \rightarrow Y_{ij}$, then $I(X_j; Y_{ij} | Y_{ej}) = 0$. In either case we can write $I(X_j; Y_{ij} | Y_{ej}) = \{I(X_j; Y_{ij}) - I(X_j; Y_{ej})\}^+$. Substituting this in (3) we have

$$R_{K,M}^{+, \text{common}} \leq \max_{p(X_1)p(X_2)\dots p(X_M)} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M \{I(X_j; Y_{ij}) - I(X_j; Y_{ej})\}^+, \quad (25)$$

which coincides with our achievable rate in (4) when we choose $U_j = X_j$.

As a special case of Theorem 1, we have the following corollary for the case of one receiver and one eavesdropper.

Corollary 1 (Single User case): Consider the reversely degraded parallel channels in Definition 2 with one receiver and one eavesdropper. The secrecy capacity is given by

$$C_{1,M} = \max_{p(X_1)p(X_2)\dots p(X_M)} \sum_{j=1}^M I(X_j; Y_j | Y_{ej}). \quad (26)$$

³It is important that the codewords be independently selected. If they are not, say the same codeword is repeated on each channel, the eavesdropper equivocation can be significantly reduced by combining the channel outputs.

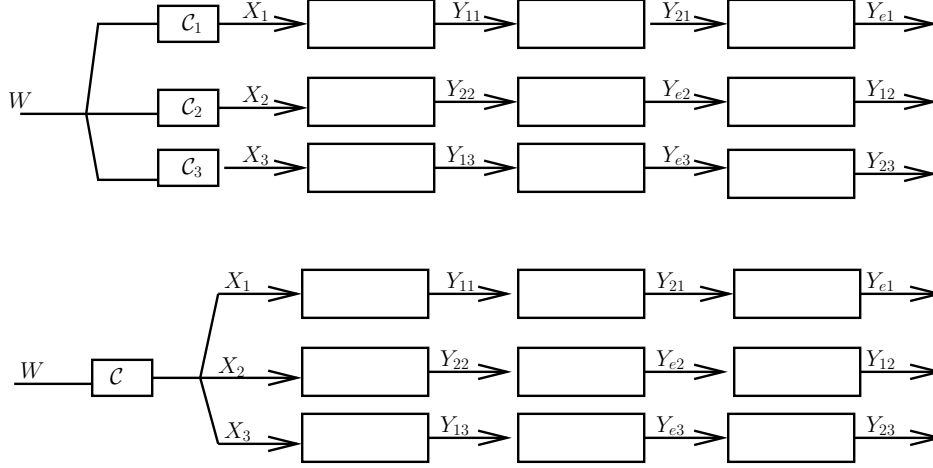


Fig. 3. Two coding schemes for common message transmission on proposed channels. The top figure shows the scheme proposed in Theorem 1. It achieves the common message capacity. In this scheme we use independent codebooks on each parallel channel. This allows us to separately bin on each channel. The lower figure shows the scheme that uses a single codebook. While this scheme is optimal when there is no eavesdropper [6], it is suboptimal in our setup. This drawback of this scheme is that because of the single codebook, one cannot separately bin for each channel.

Remark 4: The single user result admits a simple coding scheme. Split the message W into M sub-messages W_1, W_2, \dots, W_M and independently encode and decode message W_j on channel j with a codebook of rate $R_j = I(X_j; Y_j | Y_{ej})$. With multiple receivers however, this simple scheme is limited by the worst user on each parallel channel and does not achieve the secrecy capacity.

E. Sub-optimality of a Single Codebook scheme

The capacity of common message for reversely degraded broadcast channels in Definition 2 without the secrecy constraint is [6]

$$C_{K,M}^{\text{No Secrecy}} = \max_{\prod_{j=1}^M p(X_j)} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M I(X_j; Y_{ij}). \quad (27)$$

The achievability scheme in (27) uses a *single* codebook with codewords of dimension $M \times n$. The j^{th} component of the codeword is a length n sequence sampled from an i.i.d. $p(X_j)$ distribution and is transmitted on channel j .

Our achievable scheme is different from this single codebook scheme since we use independent codebooks on each parallel channel. Note that this distinction is important in achieving the secrecy capacity in Theorem 1. The distinction between these schemes is shown in Fig. 3. An achievable rate using the single-codebook scheme in our setup is

$$R^{\text{single}} = \max_{p(X_1, X_2, \dots, X_M)} \min_{i \in \{1, 2, \dots, K\}} \{I(X_1, X_2 \dots X_K; Y_{i1}, \dots, Y_{iK}) - I(X_1, X_2 \dots X_K; Y_{e1}, \dots, Y_{eK})\}. \quad (28)$$

Note that, in general, the rate (28) is smaller than (5).⁴ The intuition behind this is that, by using an independent codebook on each parallel channel, it is possible to *separately* tune the bin size on each channel according to the degradation of the eavesdropper. Finally note that our proposed scheme also provides an alternative way to [6] to achieve the common message capacity in the absence of an eavesdropper.

⁴The two expressions coincide if, for example, the eavesdropper is degraded with respect to all the receivers on all the channels, i.e., $X_j \rightarrow Y_{ij} \rightarrow Y_{ej}$ for every $1 \leq i \leq K$ and $1 \leq j \leq M$.

F. Gaussian Channels

We consider the Gaussian channel model where

$$\begin{aligned} Y_{ij} &= X_j + Z_{ij} \\ Y_{ej} &= X_j + Z_{ej}, \end{aligned} \quad (29)$$

with $Z_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$ and $Z_{ej} \sim \mathcal{N}(0, \sigma_{ej}^2)$. All these noise variables are assumed independent. We also impose an average power constraint $E[\sum_{j=1}^M X_j^2] \leq P$.

Corollary 2: The common-message-secrecy-capacity for the Gaussian parallel broadcast channel in (29) is

$$C_{K,M}^{\text{common,Gaussian}} = \max_{(P_1, P_2, \dots, P_M) \in \mathcal{F}} \min_{1 \leq i \leq K} \sum_{j=1}^M \left\{ \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{ij}^2} \right) - \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{ej}^2} \right) \right\}^+, \quad (30)$$

where \mathcal{F} is the set of all feasible power allocations that satisfy $\sum_{j=1}^M P_j \leq P$.

To prove Corollary 2, first observe that the channel in (29) has the same capacity as the corresponding reversely degraded broadcast channel (see Fact 1) given by the following model: on channel j , let $\pi_j(1), \dots, \pi_j(K+1)$ denote set of intended receivers and eavesdropper ordered from the strongest to the weakest. For each $0 \leq k \leq K$, the channel for user $\pi_j(k+1)$ is $\hat{Y}_{\pi_j(k+1)j} = \hat{Y}_{\pi_j(k)j} + \hat{Z}_{kj}$ with $Y_{\pi_j(0)j} \triangleq X_j$ and $\sigma_{\pi_j(0)j}^2 \triangleq 0$. The noise random variables $\hat{Z}_{kj} \sim \mathcal{N}(0, \sigma_{\pi_j(k+1)j}^2 - \sigma_{\pi_j(k)j}^2)$ are independent.

Since $I(X_j; \hat{Y}_{ij} | \hat{Y}_{ej})$ is a continuous and concave function in $p(X)$ (see Fact 2), we use discretization arguments (see, e.g., Ch. 7 in [8]) to extend Theorem 1 to the Gaussian case

$$C_{K,M}^{\text{common}}(P) = \max_{\substack{\prod_{j=1}^M p(X_j), \\ E[\sum_{j=1}^M X_j^2] \leq P}} \min_{i \in \{1, 2, \dots, K\}} \sum_{j=1}^M I(X_j; \hat{Y}_{ij} | \hat{Y}_{ej}). \quad (31)$$

Now observe that $\max_{p(X_j), E[X_j^2] \leq P_j} I(X_j; \hat{Y}_{ij} | \hat{Y}_{ej})$ denotes the capacity of a Gaussian wiretap channel [13]. Accordingly we have

$$\max_{p(X_j), E[X_j^2] \leq P_j} I(X_j; \hat{Y}_{ij} | \hat{Y}_{ej}) = \left\{ \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{ij}^2} \right) - \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{ej}^2} \right) \right\}^+. \quad (32)$$

One then deduces (30).

IV. PARALLEL CHANNELS: INDEPENDENT MESSAGES

We consider the case of M parallel channels, one eavesdropper and K receivers, each interested in an independent message. Each such message must be protected from the eavesdropper. We now define the achievable rate for the case of independent messages.

Definition 4 (Length n Code): A $(2^{nR_1}, 2^{nR_2}, \dots, 2^{nR_K}, n)$ code for the product broadcast wiretap channel in Definition 1 consists of a mapping $\omega_n : \mathcal{W}_1 \times \mathcal{W}_2 \times \dots \times \mathcal{W}_K \rightarrow \underbrace{\mathcal{X}^n \times \mathcal{X}^n \dots \mathcal{X}^n}_{M \text{ times}}$ from the messages

of the K users to the M channel inputs and K decoding functions $\phi_{i,n} : \underbrace{\mathcal{Y}^n \times \mathcal{Y}^n \times \dots \times \mathcal{Y}^n}_{M \text{ times}} \rightarrow \mathcal{W}_i$, one

at each intended receiver. We denote the message estimate at decoder i by \hat{W}_i . A perfect-secrecy-rate tuple (R_1, R_2, \dots, R_K) is achievable if, for every $\varepsilon > 0$, there is a length n code such that $\Pr(W_i \neq \hat{W}_i) \leq \varepsilon$ for all $i = 1, 2, \dots, K$, and such that the following condition is satisfied

$$\frac{1}{n} H(W_i | W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_K, Y_{e1}^n, \dots, Y_{eM}^n) \geq \frac{1}{n} H(W) - \varepsilon, \quad i = 1, 2, \dots, M. \quad (33)$$

The secrecy-sum-capacity $C_{K,M}^{\text{sum}}$ is the supremum of $R_1 + R_2 + \dots + R_K$ over the achievable rate tuples (R_1, R_2, \dots, R_K) .

Remark 5: Our constraint (33) provides perfect equivocation for each message, even if all the other messages are revealed to the eavesdropper. It may be possible to increase the secrecy rate by exploiting the fact that the eavesdropper does not have access to other messages. This is not considered in the present paper.

A. Main Results

Our main result is an expression for the secrecy-sum-capacity for the reversely degraded broadcast channel in Definition 2.

Theorem 2: Let π_j denote the strongest user on channel j . The secrecy-sum-capacity for the reversely broadcast channel is given by

$$C_{K,M}^{\text{sum}} = \max_{p(X_1)p(X_2)\dots p(X_M)} \sum_{j=1}^M I(X_j; Y_{\pi_j} | Y_{ej}). \quad (34)$$

Furthermore, the expression in (34) is an upper bound on the secrecy-sum-capacity when only the intended users are reversely degraded — but the set of receivers together with the eavesdropper is not degraded.

The remainder of this section will be devoted to the proof of Theorem 2 and some discussion.

B. Proof of Upper Bound in Theorem 2

We establish the upper bound in Theorem 2. Suppose a genie provides the output of the strongest receiver, π_j , to all other receivers on each channel, i.e., on channel j the output $Y_{\pi_j}^n$ is made available to all the receivers. Because of degradation, we may assume, without loss of generality, that each receiver only observes $(Y_{\pi_1}^n, \dots, Y_{\pi_M}^n)$. Clearly, such a genie aided channel can only have a sum capacity larger than the original channel. Since all receivers are identical, to compute the sum capacity it suffices to consider the situation with one sender, one receiver, and one eavesdropper.

Lemma 4: The secrecy-sum-capacity in Theorem 2 is upper bounded by the secrecy capacity of the genie aided channel, i.e., $C_{K,M}^{\text{sum}} \leq C^{\text{GenieAided}}$.

Proof: Suppose that a secrecy rate point (R_1, R_2, \dots, R_K) is achievable for the K user channel in Theorem 2 and let the messages be denoted as (W_1, W_2, \dots, W_K) . This implies that, for any $\varepsilon > 0$ and n large enough, there is a length n code such that $\Pr(\hat{W}_i \neq W_i) \leq \varepsilon$ for $i = 1, 2, \dots, K$, and such that

$$\frac{1}{n} H(W_i | W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_K, Y_{e1}^n, Y_{e2}^n, \dots, Y_{eM}^n) \geq R_i - \varepsilon. \quad (35)$$

We now show that a rate of $(\sum_{i=1}^K R_i, \underbrace{0, \dots, 0}_{K-1})$ is achievable on the genie aided channel. First, note that any message that is correctly decoded on the original channel is also correctly decoded by user 1 on the genie aided channel. It remains to bound the equivocation on the genie aided channel when the message to receiver 1 is $W = (W_1, W_2, \dots, W_K)$. We have

$$\begin{aligned} \frac{1}{n} H(W | Y_{e1}^n, Y_{e2}^n, \dots, Y_{eM}^n) &= \frac{1}{n} H(W_1, W_2, \dots, W_K | Y_{e1}^n, Y_{e2}^n, \dots, Y_{eM}^n) \\ &\geq \sum_{i=1}^K \frac{1}{n} H(W_i | W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_K, Y_{e1}^n, Y_{e2}^n, \dots, Y_{eM}^n) \\ &\geq \sum_{i=1}^K R_i - K\varepsilon \end{aligned}$$

where the last step follows from (35). Since ε is arbitrary, this establishes the claim. ■

Lemma 5: The secrecy capacity of the genie aided channel is

$$C^{\text{GenieAided}} = \max_{p(X_1)p(X_2)\dots p(X_M)} \sum_{j=1}^M I(X_j; Y_{\pi_j} | Y_{e_j}). \quad (36)$$

Proof: Since all receivers are identical on the genie aided channel, this Lemma is a direct consequence of Corollary 1 when specialized to the case of $K = 1$ receiver. ■

Remark 6: The upper bound continues to hold even if the eavesdroppers channel is not ordered with respect to the intended receivers. In general, following Lemma 1, the upper bound can be tightened by considering, for all $1 \leq j \leq M$, the worst joint distribution $p'(Y_{\pi_j}, Y_{e_j} | X_j)$ among all joint distributions with the same marginal distribution as $p(Y_{\pi_j} | X_j)$ and $p(Y_{e_j} | X_j)$, yielding

$$C_{K,M}^{\text{sum}} \leq \min_{\prod_{j=1}^M p'(Y_{\pi_j}, Y_{e_j} | X_j)} \max_{\prod_{j=1}^M p(X_j)} \sum_{j=1}^M I(X_j; Y_{\pi_j} | Y_{e_j}). \quad (37)$$

C. Achievability Scheme

Our achievability scheme for Theorem 2 requires the receivers and the eavesdropper to be reversely degraded. We only send information intended to the strongest user, i.e., only user π_j on channel j can decode. It follows from the result of the wiretap channel [22] that a rate of $R_j = \max_{p(X_j)} I(X_j; Y_{\pi_j} | Y_{e_j})$ is achievable on channel j . Accordingly the total sum rate of $\sum_j R_j$ is achievable which is the capacity expression.

Remark 7: The “opportunistic transmission” strategy in Theorem 2 has been previously observed in the absence of an eavesdropper [14], [19] in the context of fading channels. Hence our result states that the optimality of opportunistic transmission also holds in the presence of an eavesdropper. Our converse technique, when applied to the case of no eavesdropper, also provides a simpler argument for the optimality of opportunistic transmission studied in [14], [19].

D. Gaussian Channels

Theorem 2 can be extended to the case of Gaussian parallel channels. Let $\sigma_{\pi_j}^2$ denote the noise variance of the strongest user on channel j . Then the secrecy-sum-capacity is given by

$$C_{K,M}^{\text{sum,Gaussian}}(P) = \max_{(P_1, P_2, \dots, P_M)} \sum_{j=1}^M \left\{ \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{\pi_j}^2} \right) - \frac{1}{2} \log \left(1 + \frac{P_j}{\sigma_{e_j}^2} \right) \right\}^+ \quad (38)$$

where the maximization is over all power allocations satisfying $\sum_{j=1}^M P_j \leq P$. The achievability follows by using independent Gaussian wiretap codebooks on each channel and only considering the strongest user on each channel. For the upper bound we have to show that Gaussian inputs are optimal in the capacity expression in Theorem 2. The justifications are the same as in the common message case in Section III-F.

V. FADING CHANNELS

The second part of this paper considers wireless fading channels. The case of one receiver and one eavesdropper has been recently studied in a number of recent works [1], [10], [12], [15]–[17]. The proposed schemes adapt the transmission power and/or rate depending on the instantaneous channel conditions. The time varying nature of the fading channel enables secure transmission even when the eavesdropper has an average channel stronger than that of the intended receiver.

To the best of our knowledge, the above works do not consider secure transmission to *multiple* receivers in a wireless fading environment. We first consider the case when a common message has to be delivered to a set of intended receivers. Next, we consider the case when each receiver obtains an independent message. For this setting, we present a scheme based on multiuser diversity that achieves the sum capacity in the limit of a large number of receivers.

A. Channel Model

A block fading channel model for a system with one sender, K receivers, and one eavesdropper is of the form

$$\mathbf{y}_i(t) = h_i(t)\mathbf{x}(t) + \mathbf{z}_i(t), \quad i \in \{1, 2, \dots, K, e\}, \quad t \in \{1, 2, \dots, n\} \quad (39)$$

where i denotes the index of the receiver and t denotes the time index. The vectors $\mathbf{z}_i, \mathbf{x}, \mathbf{y}_i$ are T dimensional complex valued vectors, where T denotes the coherence period of the channel. The channel coefficients $h_i(t)$ are constant over a block of T symbols and change independently over the blocks.

In our analysis we focus only on the fast-fading scenario, i.e., $T = 1$. Using interleaved codebooks we can realize the fast-fading case even when $T > 1$. The fast-fading channel model is of the form

$$y_i(t) = h_i(t)x(t) + z_i(t), \quad i \in \{1, 2, \dots, K, e\}, \quad t \in \{1, 2, \dots, n\} \quad (40)$$

where the $h_i(t)$'s are sampled independently from $\mathcal{CN}(0, \mu_i)$ distribution and all the noise variables are sampled independently according to $\mathcal{CN}(0, 1)$. The input satisfies an average power constraint $E[|X(t)|^2] \leq P$.

Throughout, we assume the $h_i(t)$'s to be revealed to the transmitter, the K intended receivers and the eavesdropper in a causal manner. Implicitly we assume that there is an authenticated public feedback link from the receivers to the transmitter. The channel coefficients of the eavesdropper $\{h_e(t)\}_{1 \leq t \leq n}$ are only known to the eavesdropper. The transmitter and the intended receivers only have statistical knowledge of the eavesdropper's channel gains.

Remark 8: In our setup we are assuming only one eavesdropper. Note however that the equivocation term depends only on the statistics of $H_e(t)$ and not on the realization of $h_e(t)$. Accordingly the number of eavesdroppers does not matter as long as they are statistically equivalent and do not collude.

VI. FADING CHANNELS: COMMON MESSAGE

Secure multicasting refers to the case when each receiver is only interested in a common message. The transmitter exploits the channel knowledge of intended receivers to selectively broadcast the message to these receivers, while the eavesdropper remains ignorant of the message. Note that without the secrecy constraint, a non adaptive scheme such as the one that does a flat power allocation with no transmitter CSI, appears to be not too far from the optimal. In contrast such schemes reveal the message to an eavesdropper with a channel statistically equivalent to some intended receiver.

Perhaps, an interesting question is the scaling of the secrecy capacity with the number of receivers. Does the capacity decay to zero with the number of receivers? Note that the scheme that consists in sending information only when all the users have a strong channel performs poorly. Since the channel gains across the users are independent, the achievable rate decays to zero exponentially in the number of users.

An obvious upper bound is the secrecy capacity with a single receiver. Accordingly, the best we hope for is that the common message secrecy-capacity is a constant, independent of the number of intended receivers. In what follows, we present a coding scheme whose achievable rate is also a constant, independent of the number of intended receivers. While our proposed scheme provides optimal scaling, the precise value of the constant remains an open problem.

We now provide a formal definition of the common-message-secrecy-capacity.

Definition 5: A $(n, 2^{nR})$ code for the channel consists of an encoding function which is a mapping from the message $w \in \{1, 2, \dots, 2^{nR}\}$ to transmitted symbols $x(t) = \omega_t(w; h_1^t, h_2^t, \dots, h_K^t)$ for $t = 1, 2, \dots, n$, and a decoding function at each receiver $\hat{W}_i = \phi_i(y_i^n; h_1^n, h_2^n, \dots, h_K^n)$ for each $i = 1, 2, \dots, K$. A rate R is achievable if, for every $\varepsilon > 0$, there exists a length n code such that $\Pr(\hat{W}_i \neq W) \leq \varepsilon$ for $i = 1, 2, \dots, K$ and such that

$$\frac{1}{n} H(W \mid H_e^n, H_1^n, H_2^n, \dots, H_K^n) \geq R - \varepsilon. \quad (41)$$

The entropy term in (41) is conditioned on H_1^n, \dots, H_K^n as the channel gains of the K receivers are assumed to be known to the eavesdropper.

A. Main Results

Our main result is an achievable rate for the common-message-secrecy-rate to K receivers.

Theorem 3: An achievable common-message-secrecy-rate for the channel model (40) is given by

$$R^{\text{common}}(P) = \min_{1 \leq i \leq K} E [\{\log(1 + |H_i|^2 P) - E[\log(1 + |H_e|^2 P)]\}^+] \quad (42)$$

If all the users are i.i.d. Rayleigh faded with $E[|H_i|^2] = 1$, the following can be readily verified

$$\lim_{P \rightarrow \infty} R^{\text{common}}(P) = 0.7089 \text{ bits/symbol} \quad (43)$$

Note that the achievable rate in (42) and (43) does not depend on the number of receivers. Accordingly, we do not subscript the rate by K . That the capacity does not decay with the number of receivers is the best scaling of the capacity with the number of receivers that one can expect.

The “interesting” part of our achievability rate (42) is the $\{\cdot\}^+$ inside the expectation. This is essentially a consequence of the multiple codebook scheme we presented for the parallel channel case in Section III.

Our approach to establish the achievability of (42) is to decompose the fading channel into a set of parallel channels and invoke Lemma 2 for the *probabilistic* extension of the parallel broadcast channel.

B. Achievability Scheme

First we consider the following *probabilistic extension* of the parallel broadcast channel [14]: At each time, only one of the parallel channel operates and channel j is selected with a probability p_j , independent of all other times. Also suppose that there is a total power constraint P on the input. A straightforward extension of Lemma 2 provides the following achievable rate

$$R_{K,M}^{\text{common}}(P) \triangleq \max_{i \in \{1,2,\dots,K\}} \min \sum_{j=1}^M p_j \{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+, \quad (44)$$

where U_1, U_2, \dots, U_M are auxiliary random variables and the maximum is over the product distribution $p(U_1)p(U_2)\dots p(U_M)$ and the stochastic mappings $X_j = f_j(U_j)$ that satisfy $\sum_{j=1}^M p_j E[X_j^2] \leq P$.

Next, we map the fading channel (40) into a set of parallel channels and invoke the achievable rate (44). However, we need to resolve the technicality in that the fading channel has continuous valued fading coefficients, while the rate expression in (44) is only for a finite number of parallel channels. Following [9], our approach is to *discretize* the continuous valued coefficients and thus create parallel channels, one for each quantized state. The number of parallel channels increases as the quantization becomes finer. In what follows we only quantize the magnitude of the fading coefficients. The receiver can always rotate the phase, so it plays no part.

We quantize the channel gains into one of the q values: $A_1 = 0 < A_2 < \dots < A_{q+1} = \infty$. Receiver i is in state $l \in \{1, 2, \dots, q\}$ at time t if $A_l \leq |H_i(t)|^2 < A_{l+1}$. When in state l , the receiver's channel gain is pessimistically discretized to $\sqrt{A_l}$. Since there are K independent users, there are a total of $M = q^K$ possible super-states, which we number as S_1, S_2, \dots, S_M . Denote the quantized gain of user i in S_j by the double subscript S_{ij} . Let $p(S_j)$ denote the probability of state S_j . Also let $p_i(A_l)$ be the probability that a user i is in state l i.e., $p_i(A_l) = \sum_{k=1: S_{ik}=A_l}^M p(S_k)$. In super-state S_j , the channel of user i and the eavesdropper are

$$\begin{aligned} y_{ij}(t) &= \sqrt{S_{ij}}x(t) + z_i(t), \\ y_{el}(t) &= H_e(t)x(t) + z_e(t). \end{aligned}$$

By selecting $U_j \sim \mathcal{CN}(0, P)$ and $X_j = U_j$, the argument in the summation in (44) (with the eavesdropper output (Y_{el}, H_e)) is

$$\begin{aligned} \{I(U_j; Y_{ij}) - I(U_j; Y_{ej}, H_e)\}^+ &= \{I(X_j; Y_{ij}) - I(X_j; Y_{ej}, H_e)\}^+ \\ &= \{I(X_j; \sqrt{S_{ij}}X_j + Z_i) - I(X_j; H_e X_j + Z_e, H_e)\}^+ \\ &= \{\log(1 + S_{ij}P) - E[\log(1 + |H_e|^2 P)]\}^+. \end{aligned}$$

Substituting in (44), we have that the following rate is achievable

$$R_Q^{\text{common}}(P) = \min_{1 \leq i \leq K} \sum_{j=1}^M p(S_j) \{\log(1 + S_{ij}P) - E[\log(1 + |H_e|^2 P)]\}^+ \quad (45)$$

$$= \min_{1 \leq i \leq K} \sum_{l=1}^q p_i(A_l) \{\log(1 + A_l P) - E[\log(1 + |H_e|^2 P)]\}^+, \quad (46)$$

where the second equality follows from rewriting the summation over the states of each individual user. As $q \rightarrow \infty$, the above sum converges to

$$\min_{1 \leq i \leq K} \int_0^\infty \{\log(1 + xP) - E[\log(1 + |H_e|^2 P)]\}^+ p_i(x) dx \quad (47)$$

$$= \min_{1 \leq i \leq K} E [\{\log(1 + |H_i|^2 P) - E[\log(1 + |H_e|^2 P)]\}^+], \quad (48)$$

yielding (42).

Remark 9: The scheme presented above requires q^K codebooks, where q is the number of quantization bins. A different decomposition which requires only 2^K codebooks and provides the same achievable rate is presented in Appendix III. This scheme can be implemented in practice with an outer erasure code and an inner wiretap code, as discussed in [12].

VII. FADING CHANNELS: INDEPENDENT MESSAGES

We consider the case where each receiver wants an independent message. We will only focus on the sum rate of the system. This scenario has been widely studied in conventional systems (i.e., without a secrecy constraint) where the transmitter CSI provides dramatic gains (see e.g., [14], [19]). An “opportunistic scheme” that selects the user with the largest instantaneous gain maximizes the sum-rate of the system. The results in this section can be interpreted as an extension of opportunistic transmission in the presence of eavesdroppers.

Definition 6: A $(n, 2^{nR_1}, \dots, 2^{nR_K})$ code consists of an encoding function from the messages w_1, \dots, w_K with $w_i \in \{1, 2, \dots, 2^{nR_i}\}$ to transmitted symbols $x(t) = \omega_t(w_1, w_2, \dots, w_K; h_1^t, h_2^t, \dots, h_K^t)$ for $t = 1, 2, \dots, n$, and a decoding function at each receiver $\hat{W}_i = \phi_i(y_i^n; h_1^n, h_2^n, \dots, h_K^n)$. A rate tuple (R_1, R_2, \dots, R_K) is achievable with perfect secrecy if, for any $\varepsilon > 0$, there exists a length n code such that, for each $i = 1, 2, \dots, K$, with W_i uniformly distributed over $\{1, 2, \dots, 2^{nR_i}\}$, we have $\Pr(\hat{W}_i \neq W_i) \leq \varepsilon$ and

$$\frac{1}{n} H \left(W_i \middle| W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_K, H_e^n, H_1^n, \dots, H_K^n \right) \geq R_i - \varepsilon. \quad (49)$$

The secrecy-sum-capacity is the supremum value of $R_1 + R_2 + \dots + R_K$ among all achievable rate tuples.

A. Main Results

In the following, let H_{\max} denote the largest instantaneous channel gain among the K users. We first upper and lower bound the secrecy-sum-capacity.

Lemma 6: For the channel model (40), the secrecy-sum-capacity is upper and lower bounded as

$$R_K^+(P) = \max_{P(H_{\max}): E[P(H_{\max})] \leq P} E [\{\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max}))\}^+] \quad (50)$$

and

$$R_K^-(P) = \max_{P(H_{\max}): E[P(H_{\max})] \leq P} E [\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max}))], \quad (51)$$

respectively, where $\{v\}^+$ denotes the $\max(0, v)$.

The difference in our lower and upper bounds in (50) and (51) is that the $\{\cdot\}^+$ operator is inside the expectation in our upper bound but not in the lower bound. Thus the “loss” with respect to the upper bound occurs whenever $|H_{\max}|^2 \leq |H_e|^2$. As the number of intended receivers grows this event happens rarely and the gap between the upper and lower bounds vanishes. Formally we have

Theorem 4: The gap between our upper bound $R_K^+(P)$ and the lower bound $R_K^-(P)$ in Lemma 6 satisfies

$$R_K^+(P) - R_K^-(P) \leq \Pr(|H_e|^2 \geq |H_{\max}|^2) E \left[\log \frac{|H_e|^2}{|H_{\max}|^2} \mid |H_e|^2 \geq |H_{\max}|^2 \right]. \quad (52)$$

The bounds coincide in the limit $K \rightarrow \infty$ when all the channel gains are sampled from $\mathcal{CN}(0, 1)$.

$$C_K^{\text{sum}}(P) = \max_{P(H_{\max}): E[P(H_{\max})] \leq P} E [\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max}))] + o(1), \quad (53)$$

where $o(1) \rightarrow 0$ as $K \rightarrow \infty$.

The result of Theorem 4 shows that opportunistic transmission in conjunction with single user Gaussian codebooks achieves the optimal sum secrecy-rate in the limit of large number of receivers.

Remark 10: To the best of our knowledge, the secrecy-sum-capacity for a finite number of receivers has not been resolved for the fast fading model (40). When the coherence period is large enough so that one can invoke random coding arguments in each period, it appears possible to extend the single receiver result in [10] to determine the secrecy-sum-capacity for finite number of users. We elaborate this connection later in the section VII-F.

Our upper and lower bounds do not coincide for a finite number of users. Nevertheless, the high SNR limit provides a convenient operating regime for numerical evaluation of the bounds.

Corollary 3: We have

$$\begin{aligned} \lim_{P \rightarrow \infty} R_K^+(P) &= E \left[\left\{ \log \frac{|H_{\max}|^2}{|H_e|^2} \right\}^+ \right] \\ \lim_{P \rightarrow \infty} R_K^-(P) &= \max_{T \geq 0} \Pr(|H_{\max}|^2 \geq T) E \left[\log \frac{|H_{\max}|^2}{|H_e|^2} \mid |H_{\max}|^2 \geq T \right]. \end{aligned} \quad (54)$$

B. Upper Bound in Lemma 6

Our proof technique is to introduce a single user genie-aided channel as in Section IV and then to upper bound this single user channel. This upper bound on the genie aided channel is closely related to an upper bound provided in [10] for the slow fading channel. We nevertheless provide a complete derivation in Appendix IV.

C. Achievability in Lemma 6

The achievability scheme combines opportunistic transmission and a Gaussian wiretap code. At each time, only the message of the user with the best instantaneous channel gain is selected for transmission.

As in Section VI-B, we quantize each receiver’s channel gain into q levels $A_1 = 0 < A_2 < \dots < A_q \leq A_{q+1} = \infty$. Since the channel gains of the K users are independent, there are a total of $M = q^K$ different super-states. These are denoted as S_1, S_2, \dots, S_M . Each of the super-states denotes one parallel channel. Note that on each parallel channel, the intended users have a Gaussian channel, while the eavesdropper has a fading channel.

Our scheme transmits an independent message on each of the M parallel channels. Let $G_j \in \{A_1, A_2, \dots, A_q\}$ denote the gain of the strongest user on channel j . We use a Gaussian codebook with power $P(G_j)$ on channel j . The achievable rate on channel j is

$$\begin{aligned} R_j &= I(U_j; Y_j) - I(U_j; Y_{ej}, H_{ej}) \\ &= \log(1 + G_j P(G_j)) - E[\log(1 + |H_e|^2 P(G_j))], \end{aligned}$$

where the second equality follows from our choice of $X_j = U_j \sim \mathcal{N}(0, P(G_j))$. The overall achievable sum rate is given by

$$\begin{aligned} R_K^-(P) &= \sum_{j=1}^M \Pr(S_j) R_j \\ &= \sum_{j=1}^M \Pr(S_j) (\log(1 + G_j P(G_j)) - E[\log(1 + |H_e|^2 P(G_j))]) \\ &= \sum_{l=1}^q \Pr(A_l) (\log(1 + A_l P(A_l)) - E[\log(1 + |H_e|^2 P(A_l))]), \end{aligned}$$

where the last equality follows by using the fact that $G_j \in \{A_1, A_2, \dots, A_q\}$ and rewriting the summation over these indices. As $q \rightarrow \infty$,

$$\begin{aligned} R_K^-(P) &= \int_0^\infty (\log(1 + aP(a)) - E[\log(1 + |H_e|^2 P(a))]) p(a) da \\ &= E [\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max}))], \end{aligned} \quad (55)$$

which establishes (51).

D. Proof of Theorem 4

Let $P^*(H_{\max})$ be the power allocation that maximizes $R_K^+(P)$ in (50). We have

$$\begin{aligned} R_K^+(P) - R_K^-(P) &\leq E [\{\log(1 + |H_{\max}|^2 P^*(H_{\max})) - \log(1 + |H_e|^2 P^*(H_{\max}))\}^+] \\ &\quad - E [\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max}))] \\ &= \Pr(|H_e|^2 \geq |H_{\max}|^2) E \left[\log \frac{1 + |H_e|^2 P^*(H_{\max})}{1 + |H_{\max}|^2 P^*(H_{\max})} \mid |H_e|^2 \geq |H_{\max}|^2 \right] \\ &\leq \Pr(|H_e|^2 \geq |H_{\max}|^2) E \left[\log \frac{|H_e|^2}{|H_{\max}|^2} \mid |H_e|^2 \geq |H_{\max}|^2 \right], \\ &\leq \frac{1}{K+1} 2 \log 2 \end{aligned}$$

where the first step follows substituting the bounds in 6, the third step follows from the fact that $\log \frac{1+|H_e|^2 a}{1+|H_{\max}|^2 a}$ is increasing in a for $|H_e|^2 \geq |H_{\max}|^2$, and where the last step follows from Lemma 8 (proved in the Appendix V) and the fact that $\Pr(|H_e|^2 \geq |H_{\max}|^2) = 1/(1+K)$, since we assumed the channel coefficients to be i.i.d.

E. Proof of Corollary 3

The upper bound follows from the simple identity, that for every $P \geq 0$,

$$\left\{ \log \frac{1 + |H_{\max}|^2 P}{1 + |H_e|^2 P} \right\}^+ \leq \left\{ \log \frac{|H_{\max}|^2}{|H_e|^2} \right\}^+. \quad (56)$$

For the lower bound, we use a two level power allocation strategy in (51). Fix a threshold $T \geq 0$ and let

$$P(H_{\max}) = \begin{cases} P_0 \triangleq \frac{P}{\Pr(|H_{\max}|^2 \geq T)} & |H_{\max}|^2 \geq T \\ 0 & \text{otherwise.} \end{cases} \quad (57)$$

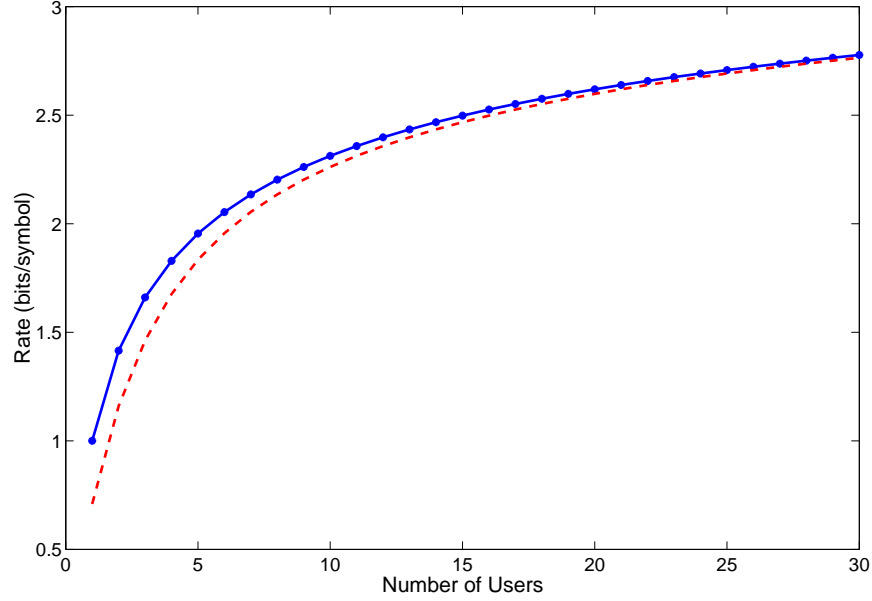


Fig. 4. Upper and Lower bounds in the High SNR limit (c.f. (54)) for the i.i.d. Rayleigh fading case. The y-axis plots the bounds in nats/symbol while the x-axis plots the number of users.

This choice gives an achievable rate of

$$R_K^-(P) = \Pr(|H_{\max}|^2 \geq T) E \left[\log \frac{1 + |H_{\max}|^2 P_0}{1 + |H_e|^2 P_0} \middle| |H_{\max}|^2 \geq T \right].$$

The argument inside the expectation is bounded by $E[\log \frac{|H_{\max}|^2}{|H_e|^2}]$ for all $P_0 > 0$. Hence by the dominated convergence theorem the limit $P \rightarrow \infty$ and the expectation can be interchanged. Accordingly we have

$$\begin{aligned} \lim_{P \rightarrow \infty} R_K^-(P) &= \Pr(|H_{\max}|^2 \geq T) E \left[\lim_{P \rightarrow \infty} \log \frac{1 + |H_{\max}|^2 P_0}{1 + |H_e|^2 P_0} \middle| |H_{\max}|^2 \geq T \right] \\ &= \Pr(|H_{\max}|^2 \geq T) E \left[\log \frac{|H_{\max}|^2}{|H_e|^2} \middle| |H_{\max}|^2 \geq T \right] \end{aligned}$$

which gives the desired result.

F. Discussion

Theorem 4 guarantees an arbitrarily small gap between upper and lower bounds on the sum-secrecy-capacity, that holds for any fixed coherence period, provided the number of users is large enough.

In [10] two schemes are presented — a *variable rate* and a *constant rate* — for the case of a single receiver in slow fading environment. Straightforward extensions of these schemes for multiple receivers reveals the following. The variable rate scheme achieves the our upper bound in (50), whereas the constant rate scheme achieve our lower bound in (51). Since these two expressions coincide as the number of receivers tends to infinity, one deduces that the gains of variable rate schemes become negligible in this limit.

Numerical Evaluation of the Upper and Lower Bounds: We plot the upper and lower bounds in the high SNR limit in (54) in Fig. 4 for the case of i.i.d. Rayleigh fading. Note that the bounds are quite close even for a moderate number of users.

Colluding Attacks: We noted earlier that any number of statistically equivalent eavesdroppers does not affect our capacity as long as they do not collude. If the eavesdroppers collude then they can combine the received signals and attempt to decode the message. The upper and lower bounds in Lemma 6 can be extended by replacing the term $|H_e|^2$ with $\|\mathbf{H}_e\|^2$, where \mathbf{H}_e is the vector of channel gains of the colluding eavesdroppers. One conclusion from these bounds is that the secrecy capacity is positive unless the colluding eavesdropper population grows as $\log K$.

VIII. CONCLUSION

A generalization of the wiretap channel to the case of parallel and fading channels with multiple receivers is considered. We established the common-message-secrecy-capacity for the case of reversely degraded parallel channels and provided upper and lower bounds for the general case. For independent messages over parallel channels, the sum-secrecy capacity has been determined. For fading channels, we examined a fast fading scenario when the transmitter knows the instantaneous channels of all the intended receivers but not of the eavesdropper. Interestingly, the common-message-secrecy-capacity does not decay to zero as the number of intended receiver grows. For the case of independent messages, it was shown that an opportunistic scheme achieves the secrecy-sum-capacity in the limit of large number of users.

The protocols investigated in this paper relied on time diversity (for the common message) and multiuser diversity (for independent messages) to enable secure communication. In situations where such forms of diversity is not available, it is of interest to develop a formulation for secure transmission, analogous to the outage formulation for slow fading channels. Secondly, the impact of multiple antennas on secure transmission is far from being clear at this stage. While multiple antennas can theoretically provide significant gains in throughput in the conventional systems, a theoretical analysis for the case of confidential messages is naturally of great interest.

APPENDIX I PROOF OF FACT 2

Let T be a binary valued random variable such that: if $T = 0$ the induced distribution on X is $p_0(X)$, i.e., $p(Y, Z, X|T = 0) = p(Y, Z|X)p_0(X)$, and if $T = 1$ the induced distribution on $p(X)$ is $p_1(X)$ i.e. $p(Y, Z, X|T = 1) = p(Y, Z|X)p_1(X)$. Note the Markov chain $T \rightarrow X \rightarrow (Y, Z)$. To establish the concavity of $I(X; Y|Z)$ in $p(X)$ it suffices to show that

$$I(X; Y|Z, T) \leq I(X; Y|Z). \quad (58)$$

The following chain of inequalities can be verified.

$$I(X; Y|Z, T) - I(X; Y|Z) = \{I(X; Y, Z|T) - I(X; Z|T)\} - \{I(X; Y, Z) - I(X; Z)\} \quad (59)$$

$$\begin{aligned} &= \{I(X; Y, Z|T) - I(X; Z|T)\} - \{I(TX; Y, Z) - I(TX; Z)\} \quad (60) \\ &= \{I(X; Y, Z|T) - I(TX; Y, Z)\} - \{I(X; Z|T) - I(TX; Z)\} \\ &= I(T; Z) - I(T; Y, Z) = -I(T; Y|Z) \leq 0. \end{aligned}$$

Equation (59) is a consequence of the chain rule for mutual information. Equation (60) follows from the fact that $T \rightarrow X \rightarrow (Y, Z)$ forms a Markov Chain, so that $I(T; Z|X) = I(T; Y, Z|X) = 0$.

APPENDIX II PROOF OF LEMMA 3

Since there are $Q_j = 2^{nR_{ej}}$ codewords per message bin $\mathcal{C}_j(W)$ and each codeword is equally likely to be selected

$$\begin{aligned} \frac{1}{n}H(U_j^n|W) &= R_{ej} \\ &= I(U_j; Y_{ej}) - \varepsilon_F, \end{aligned} \quad (61)$$

where the last equality follows from the definition of R_{ej} in (18). Since the number of codewords in each bin is less than $2^{n(I(U_j; Y_{ej}) - \varepsilon_F)}$, we can select a code that satisfies Fano's inequality

$$\frac{1}{n}H(U_j^n | W, Y_{ej}^n) \leq \gamma \triangleq \frac{1}{n} + \varepsilon_F R_{ej}. \quad (62)$$

The equivocation at the eavesdropper can be lower bounded as

$$\begin{aligned} H(W | Y_{ej}^n) &= H(W, U_j^n | Y_{ej}^n) - H(U_j^n | W, Y_{ej}^n) \\ &\geq H(U_j^n | Y_{ej}^n) - n\gamma \end{aligned} \quad (63)$$

$$\begin{aligned} &= H(U_j^n) - I(U_j^n; Y_{ej}^n) - n\gamma \\ &= H(U_j^n, W) - I(U_j^n; Y_{ej}^n) - n\gamma \end{aligned} \quad (64)$$

$$\begin{aligned} &= H(W) + H(U_j^n | W) - I(U_j^n; Y_{ej}^n) - n\gamma \\ &\geq H(W) + nI(U_j; Y_{ej}) - I(U_j^n; Y_{ej}^n) - n\gamma - n\varepsilon_F. \end{aligned} \quad (65)$$

Here (63) follows from substituting (62), (64) from the fact that W is deterministic given U_j^n and (65) follows by substituting (61). We now show that for a suitably chosen $\varepsilon' > 0$

$$I(U_j^n; Y_{ej}^n) \leq nI(U_j; Y_{ej}) + n\varepsilon'. \quad (66)$$

First note the following

$$\begin{aligned} \left| -\frac{1}{n} \log p(y_j^n) - nH(Y_j) \right| &\leq \delta, \quad \forall y_j^n \in T(Y_j) \\ \left| -\frac{1}{n} \log p(y_j^n | u_j^n) - nH(Y_j | U_j) \right| &\leq \delta, \quad \forall (y_j^n, u_j^n) \in T(Y_j, U_j). \end{aligned} \quad (67)$$

Let J be an indicator function which equals 1 if $(y_j^n, u_j^n) \in T(Y_j, U_j)$. From (67) we note that

$$I(U_j^n; Y_j^n | J = 1) \leq nI(U_j; Y_j) + 2n\delta. \quad (68)$$

Now we can upper bound $I(U_j^n; Y_j^n)$ as

$$\begin{aligned} I(U_j^n; Y_j^n) &\leq I(U_j^n; Y_j^n, J) \\ &= I(U_j^n; Y_j^n | J) + I(U_j^n; J) \\ &\leq I(U_j^n; Y_j^n | J = 1) + I(U_j^n; Y_j^n | J = 0) \Pr(J = 0) + H_b(J) \end{aligned} \quad (69)$$

$$\leq nI(U_j; Y_j) + 2n\delta + n\varepsilon \log |\mathcal{Y}| + 1, \quad (70)$$

where (69) follows from the fact that $I(U_j^n; J) \leq H_b(J)$, the binary entropy of J . The inequality (70) follows from the fact that $H_b(J) \leq 1$, $\Pr(J = 0) \leq \varepsilon$, $I(U_j^n; Y_j^n | J = 0) \leq n \log |\mathcal{Y}|$, and (68). We now select

$$\varepsilon' = 2\delta + \delta \log |\mathcal{Y}| + \frac{1}{n}.$$

Combining (65) and (70) we have

$$\begin{aligned} \frac{1}{n}I(W; Y_{ej}^n) &\leq \varepsilon' + \gamma + \varepsilon_F \\ &= 2\delta + \varepsilon |\mathcal{Y}| + \frac{2}{n} + \varepsilon_F R_{ej} + \varepsilon_F \\ &\triangleq \varepsilon'_F \end{aligned} \quad (71)$$

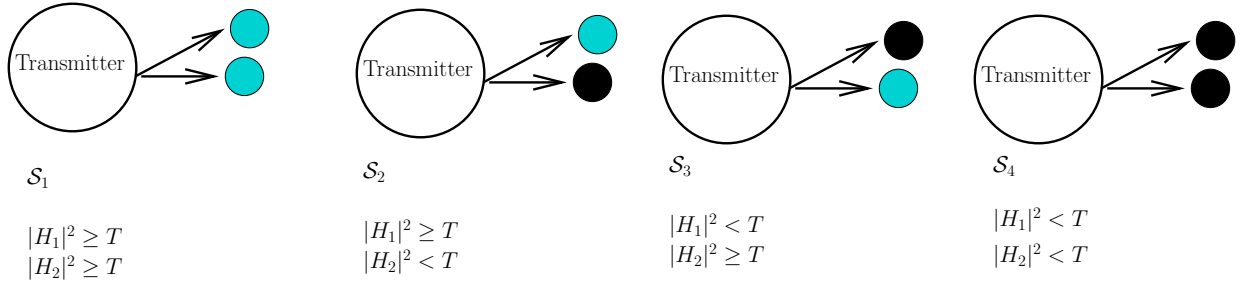


Fig. 5. Decomposition of the two user system into four states. In the first state both users have channel gains above the threshold. In the second state only user 1 has channel above the threshold while in the third state only user 2 has channel above the threshold. The fourth state both users have channels below the threshold. In any state, a user is colored dark if the channel gain is below the threshold and shaded if the channel gain is above the threshold.

APPENDIX III ALTERNATE SCHEME FOR THEOREM 3

We present an alternate scheme for Theorem 3. For simplicity we focus on the case of two receivers. The case of more than two receivers is analogous. Fix a threshold $T > 0$ and decompose the system into four states as shown in Fig. 5. . The transmission happens over a block of length n and we classify $t = 1, 2, \dots, n$ as

$$\begin{aligned}
 \mathcal{S}_1 &= \{t \in \{1, n\} \mid |h_1(t)|^2 \geq T, |h_2(t)|^2 \geq T\} \\
 \mathcal{S}_2 &= \{t \in \{1, n\} \mid |h_1(t)|^2 \geq T, |h_2(t)|^2 < T\} \\
 \mathcal{S}_3 &= \{t \in \{1, n\} \mid |h_1(t)|^2 < T, |h_2(t)|^2 \geq T\} \\
 \mathcal{S}_4 &= \{t \in \{1, n\} \mid |h_1(t)|^2 < T, |h_2(t)|^2 < T\}.
 \end{aligned} \tag{72}$$

The resulting channel is a probabilistic parallel channel with probabilities of the four channels as $p(\mathcal{S}_1) = \Pr(|H_1|^2 \geq T, |H_2|^2 \geq T)$, $p(\mathcal{S}_2) = \Pr(|H_1|^2 \geq T, |H_2|^2 < T)$, $p(\mathcal{S}_3) = \Pr(|H_1|^2 < T, |H_2|^2 \geq T)$ and $p(\mathcal{S}_4) = \Pr(|H_1|^2 < T, |H_2|^2 < T)$. Also note that with $X_j = U_j \sim \mathcal{CN}(0, P)$ in the argument of the summation in (44),

$$\{I(U_j; Y_{ij}) - I(U_j; Y_{ej})\}^+ = \begin{cases} 0, & \text{if } |H_i|^2 \leq T \text{ in } \mathcal{S}_j \\ E[\log(1 + |H_i|^2 P) - \log(1 + |H_e|^2 P) \mid |H_i|^2 \geq T], & \text{if } |H_i|^2 > T \text{ in } \mathcal{S}_j. \end{cases} \tag{73}$$

Substituting these expressions in the achievable rate expression for the probabilistic parallel channel (44) we get

$$\begin{aligned}
 R^{\text{common}}(P) &= \max_{T>0} \min_{1 \leq i \leq 2} \Pr(|H_i|^2 \geq T) E \left[\log(1 + |H_i|^2 P) - \log(1 + |H_e|^2 P) \mid |H_i|^2 \geq T \right] \\
 &= \max_{T>0} \min_{1 \leq i \leq 2} \int_T^\infty (\log(1 + xP) - E[\log(1 + |H_e|^2 P)]) p_i(x) dx \\
 &\geq \min_{1 \leq i \leq 2} \int_{T^*}^\infty (\log(1 + xP) - E[\log(1 + |H_e|^2 P)]) p_i(x) dx
 \end{aligned} \tag{74}$$

$$\begin{aligned}
 &= \min_{1 \leq i \leq 2} \int_0^\infty \{\log(1 + xP) - E[\log(1 + |H_e|^2 P)]\}^+ p_i(x) dx \\
 &= \min_{1 \leq i \leq 2} E[\{\log(1 + |H_i|^2 P) - E[\log(1 + |H_e|^2 P)]\}^+],
 \end{aligned} \tag{75}$$

where T^* in (74) is the solution to $\log(1 + xP) - E[\log(1 + |H_e|^2 P)] = 0$. (The optimality of T^* follows from the fact that $p_i(x) \geq 0$ and hence the integral is maximized by keeping all terms which are positive

and discarding the negative terms, however this is not necessary to note as this is an achievable scheme.) Note that (75) coincides with the achievable rate in Theorem 3 for the case of $K = 2$ users. As remarked earlier, this scheme straightforwardly generalizes to more than two receivers. With K receivers we will have a total of 2^K states, where each state specifies the subset of users that are above the threshold T^* .

APPENDIX IV PROOF OF THE UPPER BOUND IN LEMMA 6

Consider the channel with one receiver and one eavesdropper.

$$\begin{aligned} Y(t) &= H_{\max}(t)X(t) + Z(t) \\ Y_e(t) &= H_e(t)X(t) + Z_e(t). \end{aligned} \quad (76)$$

Along the lines of Lemma 4 in Section IV-B one deduces that the sum-secrecy-capacity of the channel (40) is upper bounded by the secrecy capacity of the genie-aided-channel (76). It remains to show that an upper bound on the secrecy capacity of this channel is

$$R^+(P) = \max_{P(H_{\max}): E[P(H_{\max})] \leq P} E \left[\log(1 + |H_{\max}|^2 P(H_{\max})) - \log(1 + |H_e|^2 P(H_{\max})) \right]^+. \quad (77)$$

In what follows we will denote the eavesdropper's channel output by $\hat{Y}_e(t) = (Y_e(t), H_e(t))$ and optimistically assume that the sequence H_{\max}^n is known to the sender and the receiver non-causally. The joint distribution of the noise variables $(Z(t), Z_e(t))$ is selected to be such that if $|H_e(t)| \leq |H_{\max}(t)|$ we have $X(t) \rightarrow Y(t) \rightarrow Y_e(t)$, otherwise we have $X(t) \rightarrow Y_e(t) \rightarrow Y(t)$.

Suppose for this channel and the sequence H_{\max}^n , there is a sequence of $(n, 2^{nR})$ codes that achieve perfect secrecy in Definition 6. Following the derivation of the upper bound Theorem 2, we have

$$\begin{aligned} nR &\leq I(W; Y^n | H_{\max}^n) - I(W; \hat{Y}_e^n | H_{\max}^n) + n\varepsilon \\ &\leq I(W; Y^n, \hat{Y}_e^n | H_{\max}^n) - I(W; \hat{Y}_e^n | H_{\max}^n) + n\varepsilon \\ &= I(W; Y^n | H_{\max}^n, \hat{Y}_e^n) + n\varepsilon \\ &\leq I(X^n; Y^n | H_{\max}^n, \hat{Y}_e^n) + n\varepsilon \end{aligned} \quad (78)$$

$$\leq \sum_{t=1}^n I(X(t); Y(t) | H_{\max}(t), \hat{Y}_e(t)) + n\varepsilon \quad (79)$$

where (78) follows from the fact that $W \rightarrow (X^n, \hat{Y}_e^n) \rightarrow Y^n$ follows a Markov chain and (79) from the fact that the channel is memoryless.

Now let \mathcal{H}_n be the set of all fades that have been realized, i.e.,

$$\mathcal{H}_n = \{\gamma \mid \exists t \in [1, n], H_{\max}(t) = \gamma\}, \quad (80)$$

let N_γ denote the number of times in the interval $[0, n]$ that the channel has fade γ , and let \mathcal{S}_γ denote the time indices corresponding to a fade γ , i.e.,

$$\mathcal{S}_\gamma = \{t \mid 1 \leq t \leq n, |H_{\max}(t)|^2 = \gamma\} \quad \gamma \in \mathcal{H}_n.$$

Letting the average transmitted power at time $t \in \mathcal{S}_\gamma$ be denoted as $P_\gamma^n(t)$ we have

$$P^n(t) \triangleq E[|X(t)|^2] \quad t \in \mathcal{S}_\gamma \quad (81)$$

where the expectation is over the set of transmitted messages and any stochastic mapping used by the encoder. Let \bar{P}_γ^n denote the average power transmitted with fade level γ

$$\bar{P}_\gamma^n = \begin{cases} \frac{1}{N_\gamma} \sum_{t \in \mathcal{S}_\gamma} P^n(t), & \gamma \in \mathcal{H}_n \\ 0 & \text{otherwise} \end{cases} \quad (82)$$

We will need the following Lemma, which follows from the capacity for the Gaussian wiretap channel [13].

Lemma 7: Let (X, Y, \hat{Y}_e) be random variables such that $Y = \sqrt{\gamma}X + Z_r$ and $Y_e = \sqrt{\mu}X + Z_r$. Suppose that $Z_r \sim \mathcal{CN}(0, 1)$ and $Z_e \sim \mathcal{CN}(0, 1)$ and that the joint distribution of (Z_r, Z_e) satisfies $X \rightarrow Y \rightarrow Y_e$ if $|\mu| \leq |\gamma|$ and $X \rightarrow Y_e \rightarrow Y$ otherwise. Then we have

$$\max_{p(X), E[|X|^2] \leq \bar{P}} I(X; Y|Y_e) = \log(1 + \gamma\bar{P}) - \log(1 + \min(\gamma, \mu)\bar{P}). \quad (83)$$

Now we have

$$\begin{aligned} & \sum_{t=1}^n I\left(X(t); Y(t) \middle| H_{\max}(t), \hat{Y}_e(t)\right) \\ &= \sum_{\gamma_0 \in \mathcal{H}_n} \sum_{t \in \mathcal{S}_{\gamma_0}} I\left(X(t); Y(t) \middle| |H_{\max}(t)|^2 = \gamma_0, \hat{Y}_e(t)\right) \end{aligned} \quad (84)$$

$$= \sum_{\gamma_0 \in \mathcal{H}_n} \sum_{t \in \mathcal{S}_{\gamma_0}} \left\{ \int_{\mu=0}^{\infty} I\left(X(t); Y(t) \middle| |H_{\max}(t)|^2 = \gamma_0, Y_e(t), |H_e(t)|^2 = \mu\right) p_{\mu} d\mu \right\} \quad (85)$$

$$\leq \sum_{\gamma_0 \in \mathcal{H}_n} \sum_{t \in \mathcal{S}_{\gamma_0}} \left\{ \int_{\mu=0}^{\infty} (\log(1 + \gamma_0 P^n(t)) - \log(1 + \min(\gamma_0, \mu) P^n(t))) p_{\mu} d\mu \right\} \quad (86)$$

$$\begin{aligned} & \leq \sum_{\gamma_0 \in \mathcal{H}_n} \sum_{t \in \mathcal{S}_{\gamma_0}} (\log(1 + \gamma_0 P^n(t)) - E_{\mu}[\log(1 + \min(\gamma_0, \mu) P^n(t))]) \\ & \leq \sum_{\gamma_0 \in \mathcal{H}_n} N_{\gamma_0} (\log(1 + \gamma_0 \bar{P}_{\gamma_0}^n) - E_{\mu}[\log(1 + \min(\gamma_0, \mu) \bar{P}_{\gamma_0}^n)]), \end{aligned} \quad (87)$$

where (84) follows by re-writing the summation over the set \mathcal{H}_n , (85) follows from the independence of $H_e(t)$ and $H_{\max}(t)$, (86) follows from Lemma 7, and (87) is justified from the fact that for any $\gamma_0 > 0$, the expression $\Psi_{\gamma_0}(y) = \log(1 + \gamma_0 y) - E[\log(1 + \min(\gamma_0, \mu)y)]$ is a concave function in y for $y \geq 0$, i.e.,

$$\sum_{t \in \mathcal{S}_{\gamma_0}} \Psi(P_{\gamma_0}^n(t)) \leq N_{\gamma_0} \Psi\left(\frac{1}{N_{\gamma_0}} \sum_{t \in \mathcal{S}_{\gamma_0}} P_{\gamma_0}^n(t)\right) = N_{\gamma_0} \Psi(\bar{P}_{\gamma_0}^n).$$

Combining (79) and (87) we get

$$R \leq \varepsilon + \sum_{\gamma_0 \in \mathcal{H}_n} \log(1 + \gamma_0 \bar{P}_{\gamma_0}^n) - E_{\mu}[\log(1 + \min(\gamma_0, \mu) \bar{P}_{\gamma_0}^n)] \frac{N_{\gamma_0}}{n}. \quad (88)$$

Now note that, for each n , we have $\int_{\gamma=0}^{\infty} \bar{P}_{\gamma_0}^n d\gamma \leq P$. Thus for each n , the set of points $\bar{P}_{\gamma_0}^n$, indexed by γ_0 , lie in a compact space. For this sequence of points, there exists a convergent subsequence $\bar{P}_{\gamma_0}^{n_i}$ that converges to some power allocation \bar{P}_{γ_0} as $n_i \rightarrow \infty$. Taking limit along the converging subsequence of the upper bound on the rate (88)

$$\begin{aligned} R & \leq \int_0^{\infty} \log(1 + \gamma \bar{P}_{\gamma}) - E[\log(1 + \min(\gamma, \mu) \bar{P}_{\gamma})] p_{\gamma} d\gamma \\ & = E[\log(1 + \gamma \bar{P}_{\gamma}) - \log(1 + \min(\gamma, \mu) \bar{P}_{\gamma})], \end{aligned}$$

and this completes the proof of the upper bound.

APPENDIX V

HELPER LEMMA IN THE PROOF OF THEOREM 4

Lemma 8: Let $H_1, H_2, \dots, H_K, H_e$ be i.i.d. unit mean exponentials. For $K \geq 2$, we have

$$E \left[\log \frac{|H_e|^2}{|H_{\max}|^2} \mid |H_e|^2 \geq |H_{\max}|^2 \right] \leq 2 \log 2$$

First note the following.

Fact 3 ([4]): Let $V_1, V_2, \dots, V_K, V_{K+1}$ be i.i.d. exponential random variables with mean λ and $V_{\max}(K+1)$ denotes the largest of these exponential and $V_{\max}(K)$ the second largest. The joint distribution of $(V_{\max}(K), V_{\max}(K+1))$ satisfies

$$V_{\max}(K+1) = V_{\max}(K) + Y, \quad (89)$$

where Y is an exponential random variable with mean λ and is independent of $V_{\max}(K)$

Proof: We have

$$E \left[\log \frac{|H_e|^2}{|H_{\max}|^2} \mid |H_e|^2 \geq |H_{\max}|^2 \right] = E \left[\log \frac{|H_{\max}|^2 + Y}{|H_{\max}|^2} \right] \quad (90)$$

$$\leq E \left[\frac{Y}{|H_{\max}|^2} \right] \quad (91)$$

$$= E[Y] E \left[\frac{1}{|H_{\max}|^2} \right] \quad (92)$$

$$= E \left[\frac{1}{|H_{\max}|^2} \right] \quad (93)$$

where (91) follows from the identity $\log(1+x) \leq x$ for $x > 0$, (92) follows from the independence of Y and H_{\max} , and (93) from the fact that $E[Y] = 1$. Since $|H_{\max}|^2 \geq \max(|H_1|^2, |H_2|^2)$ we obtain

$$E \left[\frac{1}{|H_{\max}|^2} \right] \leq E \left[\frac{1}{\max(|H_1|^2, |H_2|^2)} \right] \leq 2 \log 2.$$

■

REFERENCES

- [1] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. Int. Symp. Inform. Theory*, Seattle, July 2006.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, 1991.
- [3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, 1978.
- [4] H. A. David, *Order Statistics*. New York: Wiley, 1981.
- [5] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. IT-22, no. 6, pp. 644–654, 1976.
- [6] A. A. El Gamal, "Capacity of the product and sum of two un-matched broadcast channels," *Probl. Information Transmission*, pp. 3–23, 1980.
- [7] A. Fiat and M. Naor, "Broadcast encryption," in *Proceedings of the 13th annual international cryptology conference on Advances in cryptology*, Santa Barbara, CA, 1994, pp. 480–491.
- [8] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.
- [9] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
- [10] P. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inform. Theory*, submitted, Oct., 2006.
- [11] N. Jindal and A. J. Goldsmith, "Optimal power allocation for parallel broadcast channels with independent and common information," in *Proc. Int. Symp. Inform. Theory*, June 2004.
- [12] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting with multiuser diversity," in *Proc. 44th Allerton Conf. on Communication, Control and Computing*, 2006.
- [13] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inform. Theory*, vol. 24, pp. 451–456, 1978.
- [14] L. Li and A. J. Goldsmith, "Optimal resource allocation for fading broadcast channels- part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1083–1102, Mar. 2001.

- [15] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proc. 44th Allerton Conf. on Communication, Control and Computing*, 2006.
- [16] Y. Liang and H. V. Poor, "Secure communication over fading channels," in *Proc. 44th Allerton Conf. on Communication, Control and Computing*, 2006.
- [17] R. Negi and S. Goel, "Secret communication using artificial noise," in *Proc. IEEE Vehicular Tech. Conf.*, 2005.
- [18] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.
- [19] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," unpublished, 1999.
- [20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [21] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. IT-43, no. 2, pp. 712–14, 1997.
- [22] A. D. Wyner, "The Wiretap Channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–87, 1975.
- [23] H. Yamamoto, "A coding theorem for secret sharing communication systems with two Gaussian wiretap channels," *IEEE Trans. Inform. Theory*, vol. 37, pp. 634–638, May 1991.